# Impact of Pesticide Exposure Misclassification on Estimates of Relative Risks in the Agricultural Health Study

**Aaron Blair**[1,2], **Kent Thomas**[3], **Joseph Coble**[4], **Dale P. Sandler**[5], **Cynthia J. Hines**[6], **Charles F. Lynch**[7], **Charles Knott**[8], **Mark P. Purdue**[1], **Shelia Hoar Zahm**[1], **Michael C.R. Alavanja**[1], **Mustafa Dosemeci**[1], **Freya Kamel**[5], **Jane A. Hoppin**[5], **Laura Beane Freeman**[1], and **Jay H. Lubin**[1]
Aaron Blair: blaira@mail.nih.gov

[1]Division of Cancer Epidemiology and Genetics, National Cancer Institute

[2]National Cancer Institute, EPS Room 8008, Bethesda, MD 20892, Phone: 301-496-9093, Fax: 301-402-1819

[3]U.S. Environmental Protection Agency, Research Triangle Park, NC

[4]Annapolis, MD

[5]National Institute of Environmental Health Sciences, Research Triangle Park, NC

[6]National Institute for Occupational Safety and Health, Cincinnati, OH

[7]University of Iowa, Iowa City, IA

[8]Battelle, Inc., Research Triangle Park, NC

## Abstract

**Background—**The Agricultural Health Study (AHS) is a prospective study of licensed pesticide applicators (largely farmers) and their spouses in Iowa and North Carolina. We evaluate the impact of occupational pesticide exposure misclassification on relative risks using data from the cohort and the AHS Pesticide Exposure Study (AHS/PES).

**Methods—**We assessed the impact of exposure misclassification on relative risks using the range of correlation coefficients observed between measured post-application urinary levels of 2,4-dichlorophenoxyacetic acid (2,4-D) and chlorpyrifos metabolite and exposure estimates based on an algorithm from 83 AHS pesticide applications.

**Results—**The correlations between urinary levels of 2,4-D and chlorpyrifos metabolite and estimated exposure intensity scores from the expert-derived algorithm were about 0.4 for 2,4-D (n=64), 0.8 for liquid chlorpyrifos (n=4), and 0.6 for granular chlorpyrifos (n=12). Correlations of urinary levels with individual exposure determinants (e.g., kilograms of active ingredient used, duration of application, or number of acres treated) were lower and ranged from −0.36 to 0.19. These findings indicate that scores from an *a priori* expert-derived algorithm developed for the AHS were more closely related to measured urinary levels than the several individual exposure determinants evaluated here. Estimates of potential bias in relative risks observed in the AHS based on the correlations from the AHS/PES and the proportion of the AHS cohort exposed to various pesticides indicate that nondifferential misclassification of exposure using the algorithm would bias some estimates toward the null, but less than the misclassification associated with individual exposure determinants.

**Conclusions—**Based on these correlations and the proportion of the AHS cohort exposed to various pesticides, the potential bias in relative risks from nondifferential exposure misclassification is reduced when exposure estimates are based on an expert algorithm compared to estimates based on separate individual exposure determinants often used in epidemiologic

studies. Although correlations between algorithm scores and urinary levels were quite good (i.e., correlations between 0.4 and 0.8), exposure misclassification would still bias relative risk estimates in the AHS towards the null and diminish study power.

## Introduction

Exposure misclassification can limit the validity and precision of epidemiologic studies and diminish power to detect associations. The theory and mechanics of misclassification are well described[1–3] and the impact of exposure misclassification on relative risk estimates can be large.[4,5] In the AHS, as in many epidemiologic studies, there is no "gold standard" for exposure. In these cases, it is useful to relate estimates of exposure with actual measurements of current exposures (even if only at a single point in time) to provide an indication of the degree of exposure misclassification associated with surrogate indicators for exposures. Information from such methodologic efforts is of considerable assistance in the interpretation of epidemiologic data.

The Agricultural Health Study (AHS) is a long-term, prospective cohort study of licensed pesticide applicators and their spouses in Iowa and North Carolina.[6] The purpose of this paper is to use information from the AHS Pesticide Exposure Study (AHS/PES),[7] which compares urinary levels of pesticides with exposure estimates based on an expert-derived algorithm[8] and with several individual exposure determinants (kg of active ingredient used, hours of mixing and application, and number of acres treated) to evaluate effects of exposure misclassification on estimates of relative risks in the AHS.

## Methods

Information on pesticide use and application procedures in the AHS was obtained by self-administered questionnaires (available at http://www.aghealth.org/questionnaires.html). Questionnaire information obtained at enrollment on pesticide use included pesticides used, application methods, mixing and applying, proportion of time personally mixed pesticides, first year of use, number of years and days per year personally applied, application method, and use of protective equipment. Information obtained on specific pesticides included ever used, mixing and application method, years used, average days per year of use, and first year of use. Monitoring information from the literature and from Pesticide Handlers Exposure Database was used to develop weights for important *a priori* exposure determinants identified from the literature, including mixing, application method, repair of application equipment, and use of personal protective equipment.[8] These weights were applied to information on pesticide use practices from AHS questionnaires to create quantitative pesticide exposure intensity scores. These scores were multiplied by the lifetime days of specific pesticide use to create intensity-weighted exposure metrics that have been used in a number of epidemiologic papers on various outcomes from this cohort (the AHS bibliography is available at: http://www.aghealth.org/.

Details of the AHS/PES monitoring effort and algorithm assessment study are provided elsewhere.[7,9] Briefly, the AHS/PES participants were individuals who had completed the AHS five-year follow-up interview between 1998 and 2003, had reported use of 2,4-D or chlorpyrifos, resided in selected counties in Iowa and North Carolina, and indicated they intended to use a product containing 2,4-D or chlorpyrifos during the upcoming season. Urine spot samples and 24-hour accumulations were collected prior to, during, and after an application of the target pesticides and analyzed for levels of 2,4-D and 3,5,6-trichloro-2-pyridinol (TCP) (a metabolite of chlorpyrifos). These pesticides were selected for the assessment study because they are important agricultural chemicals worldwide, used by many AHS participants with several different application methods, and may impact human

health.[10,11] The AHS/PES participants provided information on application practices at the time of application and, in addition, the AHS/PES monitoring team recorded application practices. Both sources of information and individual exposure determinants, were used to create exposure intensity scores using the previously developed algorithm[8], and each score was compared to post application urinary levels of 2,4-D and the chlorpyrifos metabolite (TCP) using Spearman correlation coefficients. Spearman rank order correlation values were calculated because the urinary biomarker measurements were not normally distributed and because a linear relationship between biomarker measurement and exposure intensity scores could not be assumed. In addition, the algorithm scores are not fully continuous because the algorithm variable weighting factors are combined in certain discrete combinations. The pesticide exposure section of the AHS/PES questionnaire mimicked that from the five-year followup questionnaire administered to the full cohort and included questions on determinants used in the algorithm.[8] Urinary concentrations have also been compared with several individual determinants.[7,12]

We assessed the impact of exposure misclassification on relative risks from the range of correlation coefficients (0.20, 0.40, and 0.70) observed between measured urinary levels of 2,4-D and chlorpyrifos and the algorithm scores, or individual exposure determinants. We considered nine scenerios based on proportions of applicators in the AHS reporting use of various pesticides (i.e., 20%, 40%, and 70%), a range of sensitivities that are possible with correlation coefficients of 0.20, 0.40, and 0.70, and on the range of relative risks that have been observed in the AHS are often seen in epidemiologic investigations (0.5, 1.0, 2.0, and 3.0). The calculations for relative risk attenuation based on these parameters are described in the appendix. This study was approved by the National Institutes of Health Special Studies Institutional Review Board (SSIRB), protocol number OH93-NC-N013, and also by Institutional Review Boards at the University of Iowa, Westat, Inc., RTI International, and Battelle, Inc. Informed consent was obtained from all participants prior to enrollment.

## Results

Urinary biomarker measurement results have been previously reported for 2,4-D and chlorpyrifos applicators in the AHS/PES[7,9]. Geometric mean (geometric standard deviation) values in post-application urine samples were 25 (4.1) μg/L for 2,4-D applicators and 11 (2.3) μg/L TCP for chlorpyrifos. There was considerable range among the post-application measurements (greater than 600-fold for 2,4-D applicators (1.6 – 970 μg/L) and greater than 30-fold for chlorpyrifos applicators (2.5 – 80 μg/L)). Post-application geometric mean TCP levels for chlorpyrifos applicators were over seven times higher than geometric mean levels in the U.S. adult general population in the 2001 – 2002 period[13]. Geometric mean values for 2,4-D in the U.S. general population are not available due to the preponderance of non-detect values, but post-application geometric mean 2,4-D levels for 2,4-D applicators were about 20 times greater than the 95th percentile level in the U.S. adult general population[13]. Exposure intensity algorithm scores based on questionnaires were 10.3 ± 4.6 (range 1.8 – 20) for 2,4-D applicators and 9.4 ± 2.6 (range 6.6 – 14) for chlorpyrifos applicators.[9]

Spearman correlations between post application urinary levels of 2,4 D and chlorpyrifos metabolites and estimated exposure intensity scores based on monitoring team observations of AHS/PES participant activities were 0.39 for 2,4-D, 0.80 for liquid chlorpyrifos, and 0.60 for granular chlorpyrifos (Table 1).[9,12] Results were similar using exposure intensity scores based on information from participant-completed questionnaires with correlations of 0.42 for 2,4-D, 0.80 for liquid chlorpyrifos, and 0.58 for granular chlorpyrifos. Table 2 provides Spearman correlations between urinary levels of 2,4-D or chlorpyrifos metabolite among study participants and individual determinants of pesticide exposure used in some epidemiologic studies, e.g., kg of active ingredient, hours spent mixing and applying, and

number of acres treated.[12] These correlation coefficients were quite low and none was statistically significant. The correlations for 2,4-D were all less than 0.1 and those for chlorpyrifos were 0.19 for kg of active ingredient, −0.28 for hours of use per day, and −0.36 for acres treated.

Figure 1 shows the impact of exposure misclassification on relative risks considering the correlation between urinary levels and exposure estimates noted above and relative risks in a range relevant to the published results from the AHS. Correlations between estimated exposure intensity scores and urinary levels of 0.2 or less (dotted lines) and sensitivities of 0.9 or less would depress the relative risks considerably. Some lines do not provide information across the full range of possible sensitivities because they are undefined for certain combinations of prevalence of use, sensitivity, specificity, and correlation combinations. Many relative risks are so close to the null value that a reasonable interpretation would be that no association exists. For correlations of 0.4 (dashed lines), observed relative risks for the different sensitivity and exposure misclassification categories are somewhat closer to the true relative risks than for correlations of 0.2, but they still show substantial attenuation toward the null for sensitivities of 0.9 or less. Only for correlations of 0.7 (solid lines) do the observed relative risks approach the true relative risks. For true relative risks of 1.0, misclassification described here does not bias the relative risk regardless of the proportion exposed or the magnitude of the exposure misclassification, i.e., the estimated relative risk is always 1.0 and non-differential misclassification cannot create a positive association.

## Discussion

Studies have evaluated the reliability and validity of farmers' self-reports of their pesticide application activities.[14–16] The reliability of farmers' recall of the types of pesticides used is between 60% and 80% for most pesticides.[14] Farmers can also provide considerable detail regarding their application practices, although as the questions get more detailed the reliability decreases.[14] Reliable reporting of the fact of pesticide use and application technique does not, however, provide assurance that exposure metrics and, more importantly, dose can be accurately estimated from such questionnaire data. Dose, i.e., the concentration at the target tissue, is the ultimate metric of interest in epidemiologic studies, but is largely unmeasurable.[17] Exposure and biologic factors both influence dose. Only one metabolite of chlorpyrifos (TCP) was monitored in the urine in this study and the concentration of other metabolites might also be important for health outcomes, although TCP is the major chlorpyrifos metabolite in humans. Chemical-specific biologic factors at the individual level, such as permeability of the skin and other tissues of first contact and metabolism are important, but largely unavailable for epidemiologic studies. Some information on exposure factors, such as type and condition of the equipment, use of protective equipment, type of clothing, and application rate, can be obtained by interview, but with reporting error. Estimates of pesticide exposure in the AHS were developed from an algorithm that included determinants that appeared, based on the literature, to affect exposure.[8] A concern about exposure estimates based on an algorithm is that the error associated with each determinant might multiply to something quite large and unreliable. If this was true, use of a simple, single exposure determinant might be preferable to a more complicated algorithm. Thus, an indication of the magnitude of misclassification from exposure estimates based on an algorithm derived from several determinants versus estimates based on a single determinant, e.g., acres treated, hours spent mixing and applying, or amount of active ingredient used, is essential for sound interpretation of data from epidemiologic studies and to provide guidance regarding exposure estimation efforts in future studies.[18]

Data from the recent AHS/PES methodologic study found moderate to high correlations (r=0.39 to 0.80) between measured levels in the urine and algorithm-derived estimates of pesticide exposure intensity based on information from self-reports by study participants or from observations by AHS/PES investigators during the monitoring of pesticide mixing and application activity.[9] These correlations between urinary levels and algorithm scores are similar to those reported for 2,4-D, glyphosate, and MCPA elsewhere[19–21] It is important to keep in mind that comparison of observational data and monitoring data collected at the time of application does not provide direct information on farmers' ability to recall past use of pesticides, which is critical for examining relationships between chronic diseases and pesticide exposure. Whatever the correlation is between urine measurements and a farmer's reporting of specific pesticide activities at the time of monitoring, it is likely that correlation with application activities in the past would be weaker because of increased uncertainty that occurs with the passage of time. Inclusion of frequency or duration of use of pesticides in cumulative exposure indices could introduce further misclassification that would typically lead to under-estimates of risk, as has been shown elsewhere.[22] On the other hand, it is also possible that recall of the details of pesticide use over many growing seasons might provide a better estimate of cumulative exposure over a long time period than a biologic measurement of exposure from a single application, particularly because urinary levels from non-persistent pesticide exposure reflect only recent use and are not necessarily a measure of long-term use. Several conclusions can be drawn from the evaluation of the impact of exposure misclassification on estimated relative risks in the AHS. First, the correlations between questionnaire, or observer information on pesticide use, and measured urinary levels are in the range found for other factors that are usually considered to be reliably obtained for epidemiologic studies, such as tobacco and alcohol use, diet, physical activity, and health assessments.[23–28] Second, exposure estimates from an algorithm based on several determinants thought to affect exposure are more highly correlated with measured levels of these pesticides in the urine than some specific individual determinants (i.e., kg of active ingredient used, hours of mixing and application, or number of acres treated) and would result in less attenuation of relative risks. In fact, in this example the correlations between these individual determinant measures and urinary levels of 2,4-D are so low (less than 0.1) that even if the true relative risk was 3.0, the calculated relative risk would only be about 1.1, making it very unlikely that any epidemiologic study could detect an association. The correlations between these individual determinants and urinary levels of chlorpyrifos are somewhat larger (-.36 to 0.19) than for 2,4-D (–0.09 to 0.09), but they are still considerably less than found for exposure intensity estimates based on the algorithm.[8] Third, the stronger correlations between urinary levels and algorithm exposure scores (e.g., 0.4 or 0.5) would still result in considerable attenuation of observed relative risks. For example, if the correlation between algorithm exposure intensity scores and measured urinary levels was 0.4 and the true relative risk was 3.0, the observed relative risks would be between 1.3 and 1.9 when sensitivity is in the 60 to 80% range. For a true relative risk of 2.0, the observed relative risks from correlations of 0.2 or 0.4 never rise above 1.4. For true relative risks of 0.5, correlations from 0.2 to 0.4 between exposure estimates and measurements yield estimates of relative risk between 0.7 and 0.9. All of these observed relative risks are in a range where a reasonable interpretation would be that no important association exists. In the AHS/PES exposure studies, only evaluation of chlorpyrifos in the liquid formulation had a correlation of 0.7 or greater and this may be inaccurate because the sample size was very small. The attenuation of relative risks from exposure misclassification would also reduce study power, which would necessitate larger investigations to meet study objectives.

There are additional considerations in assessing the accuracy of estimates of exposure intensities used in epidemiologic studies. First, for many chronic diseases, it is generally assumed that the critical exposure window occurs many years in the past. The correlations between estimates of exposure intensity and urinary levels in the AHS/PES[7,9] are based on

simultaneous collection of information on exposure determinants by questionnaire or observation and measurement of urinary levels of pesticides. Estimates of exposure intensity based on self-reported activities that occurred years in the past would probably be subject to greater error. Second, the correlations between algorithm scores and urinary levels varied by pesticide in each of the three recent methodologic studies[9,19–21] and the range was quite large, i.e., from r=0.12 to 0.80. Third the impact of misclassification on estimates of relative risks is influenced by the proportion of individuals exposed because this affects the sensitivity and specificity levels. For the range of exposure misclassification noted here, it appears that the proportion of the population exposed was less important than the accuracy of the exposure assessment. This conclusion, however, is based on relatively thin data and a more complete evaluation of this issue is needed.

Some cautions about these findings are warranted. The AHS/PES monitoring study provides information on farmer owner/operators and may not be relevant for other pesticide applicators. The number of measurements on chlorpyrifos is quite small and estimates are relatively unstable. The differences between urinary levels and individual determinants and algorithm scores we observed need further evaluation to see if they are generalizable to other situations. However, these data provide useful evidence regarding the reliability of the exposure metrics used in the AHS and for the interpretation of AHS findings.

We draw several conclusions from our methodologic work in the AHS. First, the accuracy of reporting of pesticide use by farmers is comparable to that for many other factors commonly assessed by questionnaire for epidemiologic studies.[23–28] Second, except in situations where exposure estimation is quite accurate (i.e., correlations of 0.70 or greater with true exposure) and true relative risks are 3.0 or more, pesticide misclassification may diminish risks estimates to such an extent that no association is obvious, which indicates false negative findings might be common. Third, it appears that an algorithm that incorporates several exposure determinants into an estimate of exposure intensity predicts urinary levels better than the individual exposure determinants considered here and would result in less attenuation of relative risk estimates. This provides some confirmation of the assumption that use of algorithms will improve exposure assessment. Finally, we note that even with the reduction in power from exposure misclassification, the AHS has identified some statistically significant links between various agricultural exposures and health outcomes.[29–35]

## Acknowledgments

## References

1. Breslow, NE.; Day, NE. The analysis of case-control studies. Vol. 1. Lyon: IARC Sci Publ No. 338; 1980. Statistical methods in cancer research.

2. Checkoway, H.; Pearce, N.; Kriebel, D. Research methods in occupational epidemiology. New York: Oxford University Press; 2004.

3. Buzas, JS.; Stefanski, LA.; Tosteson, TD. Measurement error. In: Ahrens, W.; Pigeot, I., editors. Handbook of epidemiology. Springer; New York: 2004. p. 729-765.

4. Copeland KT, Checkoway H, McMichael AJ, Holbrook RH. Bias due to misclassification in the estimation of relative risk. Am J Epidemiol. 1977; 105:488–493. [PubMed: 871121]

5. Blair A, Stewart P, Lubin JH, Forastiere F. Methodological issues regarding confounding and exposure misclassification in epidemiological studies of occupational exposures. Am J Ind Med. 2007; 50:199–207. [PubMed: 17096363]

6. Alavanja MC, Sandler DP, McMaster SB, Zahm SH, McDonnell CJ, Lynch CF, Pennybacker M, Rothman N, Dosemeci M, Bond AE, Blair A. The Agricultural Health Study. Environ Health Perspect. 1996; 104:362–9. [PubMed: 8732939]

7. Thomas KW, Dosemeci M, Hoppin JA, Sheldon LS, Croghan CW, Gordon SM, Jones ML, Reynolds SJ, Raymer JH, Akland GG, Lynch CF, Knott CE, Sandler DP, Blair AE, Alavanja MC. Urinary biomarker, dermal, and air measurement results for 2,4-D and chlorpyrifos farm applicators in the Agricultural Health Study. J Expo Sci Environ Epidemiol. 2010; 20:119–134. [PubMed: 19240759]

8. Dosemeci M, Alavanja MCR, Rowland AS, Mage D, Zahm SH, Rothman N, Lubin J, Hoppin JA, Sandler DP, Blair A. A quantitative approach for estimating exposure to pesticides in the Agricultural Health Study. I. Exposure assessment. Ann Occup Hyg. 2002; 46:245–260. [PubMed: 12074034]

9. Thomas KW, Dosemeci M, Coble JB, Hoppin JA, Sheldon LS, Chapa G, Croghan CW, Jones PA, Knott CE, Lynch CF, Sandler DP, Blair AE, Alavanja MC. Assessment of a pesticide exposure intensity algorithm in the agricultural health study. J Expos Sci Environ Epidemiol. 2010; 20(6 ): 559–569.

10. Lee WJ, Blair A, Hoppin JA, Lubin JH, Rusiecki JA, Sandler DP, Dosemeci M, Alavanja MCR. Cancer incidence among pesticide applicators exposed to chlorpyrifos in the Agricultural Health Study. J Natl Cancer Inst. 2004; 96:1781–1789. [PubMed: 15572760]

11. International Agency for Research on Cancer. Overall Evaluations of Carcinogenicity: An Updating of IARC Monographs. Vol. 1 to 42. Lyon, France: 1987. IARC Monographs on the Evaluation of the Carcinogenic Risks to Humans.

12. Thomas, K. Personal communication.

13. CDC. Fourth national report on human exposure to environmental chemicals. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention; 2009. http://www.cdc.gov/exposurereport/

14. Blair A, Tarone R, Sandler D, Lynch CF, Roland A, Wintersteen W, Dosemeci, Alavanja MCR. Reliability of reporting on lifestyle and agricultural factors by a sample of participants in the agricultural health study from Iowa. Epidemiology. 2002; 13:94–99. [PubMed: 11805592]

15. Hoppin JA, Yucel F, Dosemeci M, Sandler DP. Accuracy of self-reported pesticide use duration information from licensed pesticide applicators in the Agricultural Health Study. J Exp Anal Environ Epidemiol. 2002; 12:313–318.

16. Blair A, Zahm SH. Patterns of pesticide use among farmers: implications for epidemiologic research. Epidemiology. 1993; 4:55–62. [PubMed: 8420582]

17. Cordier, S.; Stewart, PA. Exposure assessment. In: Ahrens, W.; Pigeot, I., editors. Handbook of Epidemiology. Springer-Verlag; Berlin, Germany: 2005. p. 437-462.

18. Kromhout H, Heederick R. Effects of errors in the measurement of agricultural exposures. Scand J Work Environ Health. 2005; 31 (Suppl 1):33–38. [PubMed: 16190147]

19. Acquavella J, Alexander BH, Mandel JS, Burns CJ, Gustin C. Exposure misclassification in studies of agricultural pesticides. Epidemiology. 2006; 17:69–74. [PubMed: 16357597]

20. Coble J, Arbuckle T, Lee WJ, Alavanja M, Dosemeci M. The validation of pesticide exposure algorithm using biologic monitoring results. J Occup Environ Hyg. 2005; 2:194–201. [PubMed: 15764542]

21. Arbuckle T, Burnett R, Cole D, Teschke K, Dosemeci M, Bancej C, Zhang J. Predictors of herbicide exposure in farm applicators. Int Arch Occup Environ Health. 2002; 75:406–414. [PubMed: 12070637]

22. Burstyn I, Kim H-M, Yasui Y, Cherry NM. The virtues of a deliberately mis-specified model in demonstrating a gene-environmental interaction. Occup Environ Med. 2009; 66:374–380. [PubMed: 19017698]

23. Persson PG, Norell SE. Retrospective vs original information on cigarette smoking. Am J Epidemiol. 1989; 130:705–712. [PubMed: 2773918]

24. Kelly JP, Rosenberg L, Kaufman DW, Shapiro S. Reliability of person interview data in a hospital-based case-control study. Am J Epidemiol. 1990; 131:79–90. [PubMed: 2293756]

25. Jain M, Howe GR, Rohan T. Dietary assessment in epidemiology; comparison of a food frequency and diet history questionnaire with a 7-day food record. Am J Epidemiol. 1996; 143:953–960. [PubMed: 8610709]

26. Farrow A, Farrow SC, Little R, Golding J. ALSPAC Study Team. The repeatability of self-reported exposure after miscarriage. Int J Epidemiol. 1996; 25:797–806. [PubMed: 8921459]

27. Blair SN, Dowda M, Pate RR, Kronenfeld J, Howe HG Jr, Parker G, Blair A, Fridinger F. Reliability of long-term recall of participation in physical activity by middle-aged men and women. Am J Epidemiol. 1991; 133:266–275. [PubMed: 2000844]

28. Smith TC, Smith Bl, Jacobson IG, Corbeil TE, Ryan MAK. Reliability of standard health assessment instruments in a large, population-based cohort study. Ann Epidemiol. 2007; 17:525–532. [PubMed: 17433714]

29. Blair A, Beane Freeman L. Epidemiologic studies of cancer in agricultural populations. J Agromedicine. 2009; 14(2):125–131. [PubMed: 19437268]

30. Weichenthal S, Moase C, Chan P. A review of pesticide exposure and cancer incidence I the Agricultural Health Study cohort. Environ Health Perspect. 2010; 118:1117–1125. [PubMed: 20444670]

31. Saldana TM, Basso O, Baird DD, Hoppin JA, Weinberg C, Blair A, Alavanja MCR, Sandler DP. Pesticide exposure and hypertensive disorders during pregnancy. Environ Health Perspect. 2009; 117(9):1393–1396. [PubMed: 19750103]

32. Kamel F, Tanner CM, Umback DM, Hoppin JA, Alavanja MCR, Blair A, Comyns K, Goldman SM, Korell M, Langston JW, Ross CW, Sandler DP. Pesticide exposure and self-reported Parkinson's Disease in the Agricultural Health Study. Am J Epidemiol. 2007; 165:364–374. [PubMed: 17116648]

33. Goldner WS, Sandler DP, Yu F, Hoppin JA, Kamel F, LeVan TD. Pesticide use and thyroid disease among women in the Agricultural Health Study. Am J Epidemiol. 2010; 171:455–464. [PubMed: 20061368]

34. Hoppin JA, Umbach DM, London SJ, Henneberger PK, Kullman GJ, Coble J, Alavanja MCR, Beane Freeman LE, Sandler DP. Pesticide use and adult-onset asthma among male farmers in the Agricultural Health Study. Eur Respir J. 2009; 34:1296–1303. [PubMed: 19541724]

35. Saldana TM, Basso O, Hoppin JA, Baird DD, Knott C, Blair A, Alavanja MC, Sandler DP. Pesticide exposure and self-reported gestational diabetes mellitus in the Agricultural Health Study. Diabetes. 2007; 30:529–534.

## Appendix

The plots in Figure 1 were developed based on the following procedure. Let X represent the true exposure, where X=1 denotes exposed and X=0 denotes unexposed, and similarly let Z represent the observed exposure. Suppose r denotes the correlation coefficient for X and Z, and Sen = P(Z=1 | X=1), the sensitivity, i.e., the probability an observed exposure is a true exposure. These quantities represent relationships in the general study population. Since X and Z are binary random variables, then by definition

$$r^2 = \frac{[P(Z=1, X=1) - P(Z=1)\,P(X=1)]^2}{P(Z=1)\,P(Z=0)\,P(X=1)\,P(X=0)}$$

which can be rewritten as

$$r^2 = \frac{[Sen - P(Z=1)]^2 \, P(X=1)}{P(Z=1) \, P(Z=0) \, P(X=0)}$$

and as a quadratic equation in $P(Z=1)$,

$$P(Z=1)^2 \, [r^2 + P(X=1)(1-r^2)] - P(Z=1)[r^2 + P(X=1)(2\,Sen - r^2)] + Sen^2 P(X=1) = 0$$

that can be solved to obtain $P(Z=1)$. Since $P(Z=1) = Sen\,P(X=1) + (1-Sp)\,P(X=0)$, where $Sp$ = $P(Z=0 \mid X=0)$ is the specificity, i.e., the probability that an observed non-exposure is a true non-exposure, we can solve for $Sp$ as

$$Sp = \frac{1 - P(Z=1) - P(X=1)\,(1-Sen)}{1 - P(X=1)}$$

We assume misclassification is non-differential, which implies that $Sen$ and $Sp$ are not related to case status, that is, the same in the general population and in case subjects. Note that while $Sen$ and $Sp$ do not depend on case status, the correlation coefficient, $r$, does depend on the probability of exposure. Thus, $r$ in cases will in general not equal $r$ in the general population if the exposure factor is related to disease outcome.

For a cohort study and for disease outcome $D$, where $D=1$ denotes disease and $D=0$ denotes disease-free, the probability of disease for observed exposure $Z=1$, denoted $P(D=1 \mid Z=1)$, can be expressed as

$$\begin{aligned}
P(D=1|Z=1) &= P(D=1, X=1|Z=1) + P(D=1, X=0|Z=1) \\
&= [P(D=1, X=1, Z=1) + P(D=1, X=0, Z=1)]/P(Z=1) \\
&= [Sen\,P(D=1|X=1)\,P(X=1) + (1-Sp)\,P(D=1|X=0)\,P(X=0)]/P(Z=1) \\
&= [Sen\,RR_{true}P(X=1) + (1-Sp)\,P(X=0)]\,P(D=1|X=0)/P(Z=1)
\end{aligned}$$

where $RR_{true}$ is the true relative risk and $RR_{true} = P(D=1|X=1)/P(D=1|X=0)$. The third line follows from the assumption of non-differential misclassification, or equivalently that the observed exposure provides no additional information on disease outcome once the true exposure status is known, i.e., $P(D|X,Z) = P(D|X)$.

Following a similar process, we obtain

$$P(D=1|Z=0) = [(1-Sen)\,RR_{true}P(X=1) + Sp\,P(X=0)]\,P(D=1|X=0)/P(Z=0)$$

Thus, the observed relative risk ($RR_{obs}$) can be expressed as

$$RR_{obs} = \frac{P(D=1|Z=1)}{P(D=1|Z=0)}$$
$$= \frac{Sen\ RR_{true}\ P(X=1) + (1-Sp)\ P(X=0)}{(1-Sen)\ RR_{true}\ P(X=1) + Sp\ P(X=0)} \times \frac{P(Z=0)}{P(Z=1)}$$

For each $P(X=1)$, sensitivity, $RR_{true}$ and r, the corresponding $RR_{obs}$ for the figure is obtained by first solving the quadratic equation for $P(Z=1)$, then calculating $RR_{obs}$ from the above equation.

In a similar way, a comparable expression can be developed for true and observed relative risks, $OR_{true}$ and $OR_{obs}$, respectively, in a case-control setting, namely,

$$OR_{obs} = \frac{P(Z=1|D=1) \times P(Z=0|D=0)}{P(Z=1|D=0) \times P(Z=0|D=1)}$$
$$= \frac{Sen\ OR_{true}\ P(X=1|D=0) + (1-Sp)\ P(X=0|D=0)}{(1-Sen)\ OR_{true}\ P(X=1|D=0) + Sp\ P(X=0|D=0)} \times \frac{P(Z=0|D=0)}{P(Z=1|D=0)}$$
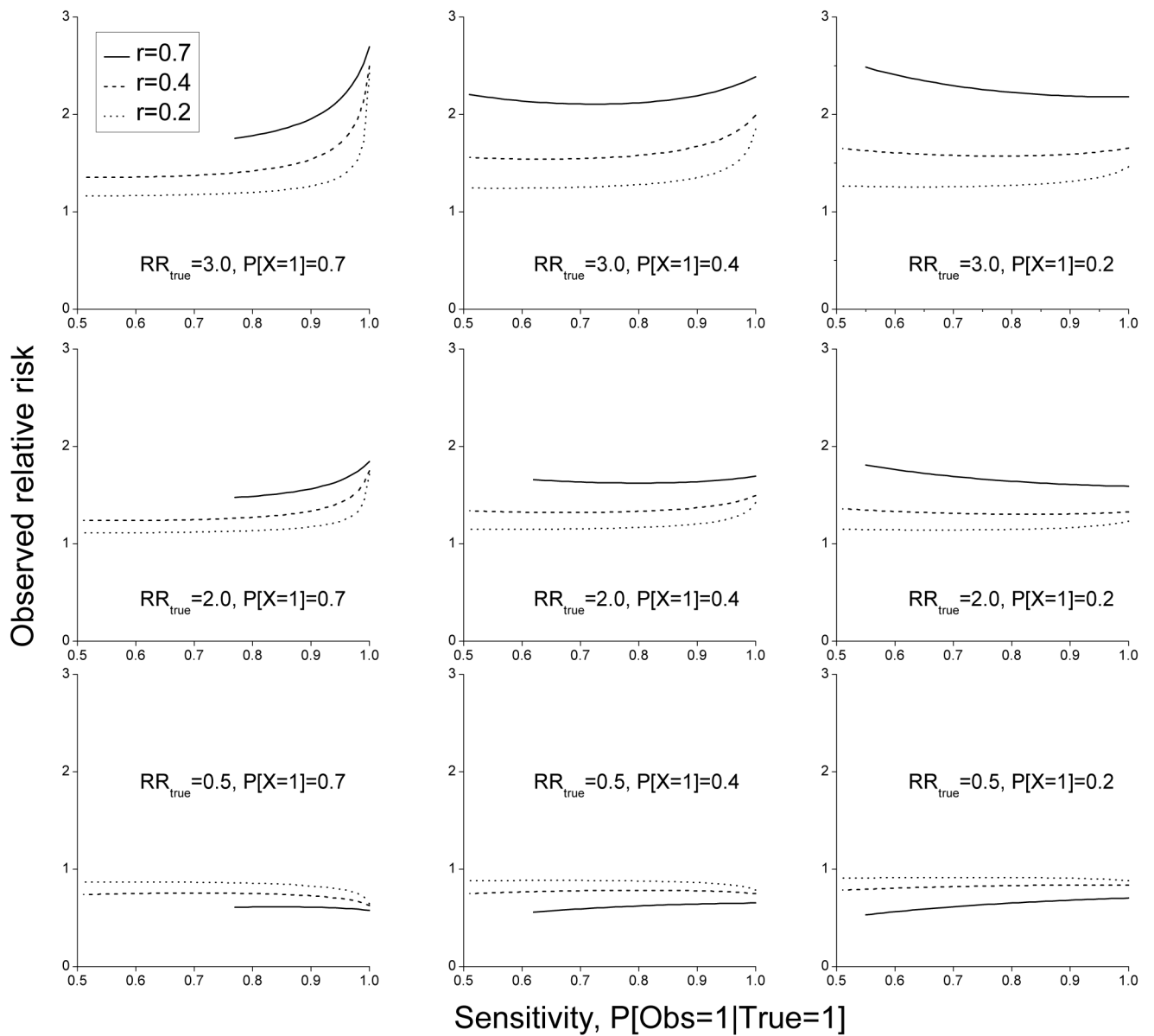
**Figure 1.**
Plots of observed relative risks based on different correlations between estimated exposure intensity scores and urinary levels (See appendix for further description of these plots).

**Table 1**

Spearman correlations between calculated pesticide exposure intensity scores and post-application urinary levels in the Agricultural Health Study Pesticide Exposure Study.[14]

| Intensity Score Source | 2,4-D (N=68) | Chlorpyrifos[+] (Liquid formulation) (N=4) | Chlorpyrifos[+] (Granular formulation) (N=12) |
|---|---|---|---|
| Observation | 0.39[**] | 0.80 | 0.60[*] |
| Questionnaire | 0.42[**] | 0.80 | 0.58[*] |

[*] 0.01<p<=0.05;

[**] p<=0.001

[+] Chlorpyrifos metabolite measured was 3,5,6-trichloro-2-pyridinol (TCP)

**Table 2**

Spearman correlations between exposure surrogates and post-application urinary levels in the Agricultural Health Study Pesticide Exposure Study (Thomas et al., Personal Communication).

| Pesticide Applied | KG Active Ingredient | Hours Mixed or Applied | Acres Treated |
|---|---|---|---|
| 2,4-D (N=63 to 68)[1] | 0.05 | 0.09 | −0.09 |
| Chlorpyrifos (N=16) | 0.19 | −0.28 | −0.36 |

[1]Number of individuals with monitoring data varied for the three determinants.